View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Journal of Phonetics 48 (2015) 76–95

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Contextual activation of Australia can affect New Zealanders' vowel productions

Kauyumari Sanchez, Jennifer Hay *, Elissa Nilson

New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

ARTICLE INFO

Article history:
Received 25 November 2013
Received in revised form
15 September 2014
Accepted 15 October 2014
Available online 13 November 2014

Keywords:
Speech production
Priming
Style-shifting
Topic
New Zealand English
Australian English

ABSTRACT

The short front vowels KIT /ɪ/, TRAP /æ/, and DRESS /ɛ/differ in their realization between speakers in New Zealand and Australian English. This paper analyses how New Zealanders produce these vowels when in an Australian-primed context. Two studies are undertaken. The first – a corpus analysis – looks at the realization of these vowels in New Zealanders' spontaneous talk about Australia. The second – an experiment – looks at the realization of these vowels in a word reading task, following the production of Australia-related lexical items. Both the experiment and the corpus analysis show differences in participant productions across Australia and non-Australia contexts. The corpus analysis shows a significant effect on the realization of the KIT and TRAP vowels, with Australian contexts associated with more Australian realizations. Both the corpus and the experiment reveal a significant interaction between speaker experience and context for DRESS. Only speakers who have ample previous experience with Australian English produce more Australian variants in an Australian context. These results highlight how different methodological approaches can provide different angles on the same question. Together, they show that subtle topic-based variation in speech production can occur. They also indicate that individual speakers' experience and beliefs can also play an important mediating role.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

This paper asks whether contextual reference to a known dialect region can shift speaker's vowel productions in the direction associated with that region. The particular population under study are speakers of New Zealand English, most of whom have had at least some exposure to Australian English. Several recent studies have demonstrated that New Zealanders seem to respond to Australia related primes in a way that affects their speech perception (Drager & Hay, 2010; Hay & Drager, 2010; Hay, Nolan, & Drager, 2006). This paper extends this line of research by investigating the effect of Australia related information on the *production* of New Zealand vowels. First, in a corpus investigation, we find that there are phonetic consequences for the production of some short front vowels when New Zealanders are talking about topics relating to Australia. This finding is important in establishing that such effects are not confined to the laboratory, as they are demonstrated here to occur in a natural conversational setting.

Second, an experiment was conducted to follow-up on the corpus results to further investigate the role of preceding Australia-related words (primes) on vowel production in an experimentally controlled manner. We find that there are phonetic consequences for the production of a word when it is immediately preceded by two words relating to Australia. In addition, the experimental work revealed that a speaker's experience with Australia plays an important role in speech shifts – an effect that is replicated when revisiting the corpus data. The results demonstrate that the mention of Australia can affect vowel productions in both experimental and relatively natural settings. Taken together, these findings also exemplify the usefulness of pursuing questions with a dual corpus-based/experimental approach.

* Corresponding author. Tel.: +64 3 364 2987x6242.
E-mail address: jen.hay@canterbury.ac.nz (J. Hay).

http://dx.doi.org/10.1016/j.wocn.2014.10.004
0095-4470 © 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1.1. New Zealand and Australian English

The regional dialects of interest in this study are New Zealand (NZE) and Australian (AE) English. This work builds on a series of past investigations on these dialects; as such it is useful to discuss the differences between these dialects. In this investigation, the short front vowels will be emphasized, as the differences between the dialects in these vowels are well documented. Although the front short vowels differ markedly, the other monopthongs of NZE and AE share similar locations in vowel space (Watson, Harrington, & Evans, 1998). While there is, of course, structured variation within each dialect with respect to the realization of these vowels (see e.g. Cox & Palethorpe, 2008; Maclagan & Hay, 2007), the overall differences observed in production between the dialects are relatively sizeable. This paper therefore concentrates on the realizations of the short front vowels: /ɪ/, /æ/, and /ɛ/, which we will refer to as the KIT, TRAP, and DRESS vowels, respectively, following Wells (1982).

Among these three vowels, the stressed KIT vowel is the most readily identifiable linguistic feature that separates NZE and AE (Watson et al., 1998). The AE variant tends to be high and front compared to the more central location of the NZE variant (Watson et al.,1998). This difference is widely known and is salient to speakers of each variety, and it is often the target of humor from speakers of both dialects. Thus, the KIT vowel serves as a salient and stereotyped feature between NZE and AE. In both dialects, unstressed KIT has a tendency to be centralized.

The TRAP and DRESS vowels are also features that distinguish NZE and AE. Both are lower and backer in AE than in NZE (Watson et al.,1998). In fact, this difference has been increasing over the last two decades, with the lowering of TRAP and DRESS in AE documented over the period 1995–2007 (Cox & Palethorpe, 2008), and the rising of the same vowels in NZE over a similar period (Maclagan & Hay, 2007). Fig. 1 illustrates the approximate relative positions of these three vowels in Australian and New Zealand English. Note that these overall general positions can also be influenced by selected phonological conditioning environments. DRESS and TRAP are merged before /l/ in NZE, for example (Hay, Drager, & Thomas, 2013), and TRAP can be raised in Australian English before nasals (Cox & Palethorpe, 2008).

In comparison, the differences in TRAP and DRESS are lesser known to NZ listeners than the difference in KIT. For example, in an open-response question, Ludwig (2007) asked NZE speakers to list the features that distinguish Australian and New Zealand English. Of her 40 respondents, over half (60%) mentioned vowels as features distinguishing AE from NZE, but few listed specific vowels. For those who did choose to list specific vowels, respondents were more likely to mention KIT (25% of respondents), than TRAP (10% of respondents), and were least likely to mentioned DRESS (5% of respondents) as distinguishing features. This suggests that the differences in NZE and AE for TRAP and DRESS are less readily identified compared to KIT.

In addition, Ludwig also tested NZE speakers' accent identification ability by using synthesized stimuli that matched typical realizations of New Zealand and Australian KIT, TRAP, and DRESS vowels. The New Zealand respondents were the most accurate in identifying the Australian KIT as Australian, with 83% of respondents correctly identifying the accent as Australian when the word contained an Australian KIT variant. Comparatively, respondents were far less accurate for TRAP. Here, performance was at chance, with 51% of the respondents correctly identifying the accent as Australian when the word contained an Australian TRAP variant. Finally, respondents were wrong more often than they were right when attempting to identify the regional origin of DRESS tokens. More respondents thought that Australian DRESS variant was more typical of an NZ accent (56%) than the AE accent (44%).

In sum, the key features that distinguish NZE and AE are KIT, TRAP, and DRESS. NZE KIT is more central than the high front AE KIT and the stereotypes are well known to speakers of both dialects. NZE TRAP and DRESS are higher than the AE variants, but are less known to speakers of both dialects. Although responses to TRAP seem to be consistent in their representation for both dialects, responses for DRESS seem to be flipped and – if anything – New Zealanders *incorrectly* associate a lowered DRESS vowel more with a New Zealand identity than an Australian one and incorrectly attribute NZE-featured items as Australian.

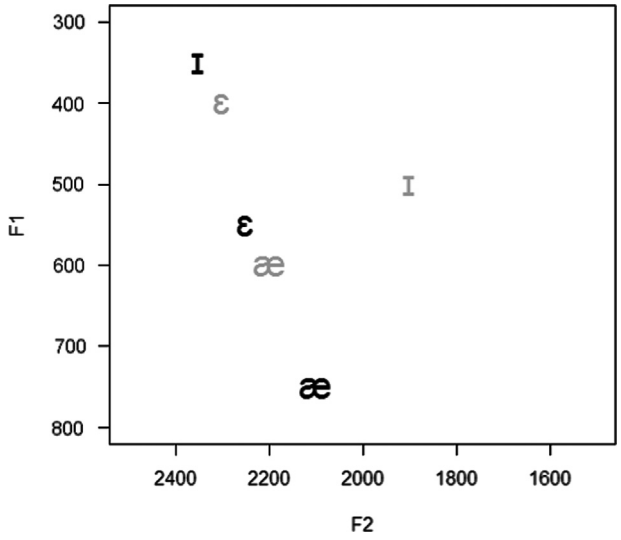

Fig. 1. Approximate relative position of DRESS, TRAP and KIT vowels in New Zealand English (gray) and Australian English (black). Approximate positions are based on Bauer, Warren, Bardsley, Kennedy and Major (2007) and Cox & Palethorpe (2008).

1.2. Addressee-based and topic-based style-shifting

Research in speech accommodation (also referred to in the literature as speech convergence and speech alignment) shows that conversation partners may spontaneously shift their speech toward their fellow interlocutor (Bell, 1984; Giles, Coupland, & Coupland, 1991). However, there is also a wealth of research that has found accommodation in socially devoid experimental contexts toward the speech of a model's recorded voice (Goldinger, 1998; Miller, Sanchez, & Rosenblum, 2010; Miller, Sanchez, & Rosenblum, in press; Namy, Nygaard, & Sauerteig, 2002; Nielsen, 2011; Sanchez, Miller, & Rosenblum, 2010; Shockley, Sabadini, & Fowler, 2004; Sanchez, 2011) or face (via lip-reading) (Miller et al., 2010; Sanchez et al., 2010; Sanchez, 2011). In addition, it has been found that a perceiver of a model's speech may be affected up to a week (insofar as it was tested) after exposure in the direction of the perceived model when simply reading text words (Goldinger & Azuma, 2004). What is more, these accommodation effects have been found across several languages and are also found in the speech of non-native talkers of a second language (e.g. Kim, Horton, & Bradlow, 2011; Lewandowski, 2012; Nguyen, Dufour, & Brunellière, 2012; Olmstead, Viswanathan, Aivar, & Manuel, 2013).

Furthermore, other laboratory studies have found that attitudes and the source of the influencing speech information can affect speech productions. For example, Babel (2010) found that New Zealanders shift their vowels (e.g. *F*1 and *F*2) towards an Australian speech style (for some vowels) when repeating recorded words said by an Australian English speaker. However, accommodation toward an Australian speech style varied across vowels, and in some cases was mediated by the subjects' implicit attitudes. In a similar vein, Babel (2012) also found that accommodation can be affected by perceptions of likeability of the model speaker. Furthermore, it has been found that the target voice does not need be attributed to an overt model, as speakers are found to shift toward voices that are simply present in the ambient environment (Delvaux & Soquet, 2007). These results indicate that one's phonetic productions can be affected by the speech of others, even speech in the environment, and that one's preconceptions or attitudes towards a speaker can also influence the effect.

Work from less experimental settings shows that individuals also shift their productions in principled ways that are not influenced by the direct interlocutor or perceived model, but rather by the general topic of conversation. Several studies suggest that individuals can display phonetic style-shifting across different conversational topics. Such results can be associated with shifting along a formality continuum within a single dialect (Labov, 1972), influenced by the identity of a person being discussed (Mendoza-Denton, Hay, & Jannedy, 2003), or by the most common interlocutor for a particular topic (Bell, 1984).

There are also many recent cases where shifting across different conversational topics and contexts is argued to be associated with different aspects of a speakers' identity (e.g. Becker, 2009; Eckert, 2000; Lawson, 2009; Love & Walker, 2013). For example, Love and Walker (2013) found that speakers would shift their pronunciations when discussing topics associated with their identity.

Love and Walker (2013) interviewed American and British born men who identified as soccer fans at a "soccer pub" in the U.S. All participants had moderate exposure to both US and British English. It was found that both American and British English speakers became more /r/-full when talking about American Football than soccer (which is associated with England, and thus predominantly /r/-less dialects). Love and Walker suggest that the variation in /r/-fullness may be due to these topics having a close relationship with the participants' identity (e.g. self-classifying as a fan of football and soccer or only soccer). However, they also raise the possibility that the effect may be more automatic and less identity-oriented than this.

In this paper, we are interested in whether New Zealanders' speech productions have a different phonetic realization when talking about Australia, compared to speech not concerning Australia. Although the countries have a close relationship, there is also a bit of a rivalry (insofar as politics and international sports are concerned) between New Zealand and Australia. Rivalry notwithstanding, it would be surprising for most New Zealanders to claim an 'Australian-oriented' identity. Nevertheless, most New Zealanders have had a reasonable amount of exposure to Australian English, via travel, the media, or personal connections.

Although the identity explanation should be explored further in the theoretical landscape of speech, the exposure explanation falls within the explanatory power of an exemplar approach. In an exemplar model, perception of speech (e.g. experiences) creates information rich memories, indexing information about the environment, the word, and information about the talker uttering the word, which may also include social information, such as the dialect of the talker (Goldinger, 1998; Pierrehumbert, 2001, 2002, 2006).

In speech perception (which will be discussed in the next section), items in memory that are related to the item being perceived (e.g. via relationship to the environment, word, or source), are activated and brought into working memory (Goldinger, 1998; Pierrehumbert, 2001). The items that most closely match the item being perceived are activated first and given more weight than less similar items in memory. Items that are given more weight serve to facilitate perception and inform production. In production one essentially *speaks out* the information activated in working memory. This process can explain the shifts in speech obtained from accommodation and topic-shifting studies.

Such models would predict automatic topic-based shifting effects, assuming encountered speech produced by Australians is indexed with information associated with 'Australia'. Thus, activation of the concept of Australia may slightly bias productions in the direction of the activated Australian memories, leading to Australian-like speech productions.

Initial support for this hypothesis may be found in Drager et al.'s (2010) experiment. In a word list reading task, Drager et al. found that New Zealanders produce different variants of the KIT vowel following exposure to a set of biased (good or bad) Australia-related facts compared to those exposed to neutral facts (not concerning Australia). Speech production shifts were influenced by attitudinal factors toward Australia. These results cannot be considered a typical example of 'topic-based' style shifting, as the participants were not in conversation; participants simply read words in a list before and after being exposed to information. Yet, this study suggests that contextual effects of New Zealanders 'talking about Australia' on phonetic production may also lead to speech shifts.

Thus, a combination of experimental studies and sociolinguistic analyses leads us to hypothesize that New Zealanders' productions of (some of) the short front vowels will be influenced by Australian conversational topics and/or by Australian lexical primes. This type of Australia-primed shifting effect has certainly been explored in some depth in the context of speech perception, as we summarize below.

1.3. Effects of contextual primes on speech perception

Past experiments have found that presenting information concerning a known dialect (via location labels or associative objects) can lead to differences in speech perception of the same speech, leading to perceptions that are congruent with the dialectical prime. For example, Niedzielski (1999, 1997) found that participants perceived different realizations of the same talker's speech depending on whether the words “Michigan” or “Canadian” were on the participant's answer sheet. In addition, Hay et al. (2006) were able to replicate Niedzielski's (1999, 1997) general findings using New Zealand and Australian English. In their examination, participants appeared to perceive the vowels of the same talker differently depending on whether the words “New Zealand” or “Australia” were on the participant's answer sheet. Interestingly, all participants (except one) responded that they believed they were listening to a New Zealander. In addition, Jannedy, Weirich, and Brunner (2011) found a related result for German listeners in Berlin. In their study, a label indicating the name of one of two neighborhoods influenced listeners’ position of German phoneme boundaries in a synthesized fricative continuum. Thus, activating an idea or concept that has connections with particular dialect may be sufficient to shift one's speech perceptions.

A further test of the power of concept activation was conducted by Hay and Drager (2010), who used items associated with members of a particular dialect instead of words. Using the auditory stimuli from Hay et al. (2006), the concepts of New Zealand (where both experiments were conducted) and Australia were introduced in the experiment via stuffed toy animals, specifically stuffed toy kiwi birds or stuffed kangaroos and koalas. These toys were present in the experiment room, but appeared to be unrelated to the experiment. Here too it was found that the same stimuli were perceived to sound differently, patterning along dialectal lines, depending on whether the items in the room were associated with New Zealand or Australia.

Hay, Warren, and Drager (2010) also review a range of studies showing that short term exposure to, or priming of, one dialect can affect the processing of another. In their studies, New Zealanders exposed to US or UK English performed differently on tasks involving vowels undergoing merger in New Zealand. The nature and magnitude of the difference was mediated by the degree to which participants were involved in the merger themselves.

Taken together, the results of these studies suggest that the activation of a known dialect (via words or items) can influence speech *perceptions* in line with the activated dialect. This perceptual tuning may be beneficial as it may serve to facilitate communication goals (Pickering & Garrod, 2004).

1.4. Interim summary

How are the automatic perceptual priming results outlined in Section 1.3 related to the natural variation individuals seem to display as they converse across topics (Section 1.2)? It seems likely that such effects are related. Work from experimental and from sociolinguistic investigations shows the degree to which individuals' production and perception can be surprisingly malleable across contexts and priming conditions. Yet, work in this area employs disparate methodologies for disparate research questions relating to different phonetic and social variables, all of which make the results from study to study difficult to compare. For example, investigations using natural speech face problems with lack of experimental control and comparability across speakers, and have challenges associated with replication. Experiments finding automatic priming effects, on the other hand, can be criticized for their lack of ecological validity. One can certainly question, for example, the degree to which results such as Hay and Drager's (2010) stuffed toys experiment reflect anything that could be observable in the ‘real world’, away from the constraints of the tightly controlled environment. Thus, the current investigation aims to connect natural and experimental research methods in the hopes of finding complementary results and an opportunity to directly compare findings.

In this paper we attempt to address whether activation of a known dialect can influence speech production. We take a two-pronged approach to this. First, we analyze productions from a corpus of spontaneous speech – comparing the productions of vowels in Australia-relevant contexts versus other contexts. The result of this analysis raises interesting questions which we pursue within a controlled experiment. The experiment shows a strong interaction between the magnitude of shift across Australia/non-Australia priming conditions, and the degree to which the speaker has previous experience of Australia and Australians. With this insight, we revisit our corpus work and investigate whether experience with Australia affects our previous results. The interplay between jointly pursued corpus and experimental work allows us to shed light on a phenomenon in a way that neither study could independently.

2. Canterbury corpus study

2.1. Method

2.1.1. Participants and materials

The Canterbury Corpus is part of the Origins of New Zealand English (ONZE) project. It consists of a sociolinguistic interview and reading of a word list, with approximately 450 speakers born from 1930 to 1984. These interviews have been completed by

undergraduate students as part of their course requirements. This corpus has been collected and added to yearly since 1994 (see Gordon, Maclagan, & Hay, 2007 for details).

Fifteen NZE speakers (eight female, seven male) between the ages of 20–59 at time of interview (mean age = 38.5) from the Canterbury Corpus were included in our analysis. These speakers were the only speakers identified as having a reasonable section of conversation where Australia was the main focus, and some period of speech where it was not. The date of interview for these speakers ranged from 1994 to 2006 (median year of interview = 1995).

2.1.2. Procedure

In order to identify relevant speakers and sections for analysis, the Canterbury Corpus was searched using the software LaBB-CAT (Fromont & Hay, 2008, 2012) for all occurrences of the word "Australia" and other Australia-relevant words, (a full list is provided in Appendix A). Search results were then combined to identify speakers that had multiple Australian-like mentions. The transcripts of speakers that mention Australia or any Australian-like words were then read and notes were taken on the content, and length of any speech mentioning Australia.

Speakers with a minimum of 30 s of reference to Australia, but excluding those who only had Australian-referring speech, were then selected for analysis. Australian-referring speech included anything that mentions time spent traveling or visiting Australia, and time spent living in Australia, though the activities described were often not particularly unique to Australia. It also included discussion of Australian-related topics such as sports teams or television speakers and programmes.

The target vowels were KIT, TRAP, and DRESS. All open class words with a target vowel containing primary stress, including multiple occurrences of the same word that occurred during the Australian speech, were selected for analysis. When the number of desirable tokens was low, both closed class and open class words with the target vowels in unstressed positions were used, provided the word was not completely prosodically unstressed. In these circumstances, there was an attempt to select a lexical match or equivalent in Australian and Non-Australian (or neutral) speech contexts. As a result of this, 13% of tokens were closed class items. Thus, while the exact composition of the target words varies across speakers, the distribution of environments is somewhat matched within each speaker, to be comparable across the 'Australian' and 'Non-Australian' samples.

Non-Australian (or neutral) speech was chosen to maximally reflect the speaker's 'standard' New Zealand English speech. Thus, speech referring to other places (e.g. America, Canada, England) was avoided. The Non-Australian (or neutral) speech was selected at the furthest point (temporally) from the Australian speech and was approximately matched in quantity to the Australian speech.

A small number of target words were explicitly associated with Australia. These were mainly place-names (e.g. Sydney). These cases were always in the 'Australian' context. We tested our models with these words included and excluded. There were no substantive differences in the models. In order to be conservative, we report here the models which exclude these words. Only 'non-Australian' lexical items are included in the analysis reported below.

2.1.3. Acoustic measurements

The first and second formants (*F*1 and *F*2, respectively) of the selected tokens' vowels were measured using Praat (Boersma & Weenink, 2008). The measurements were taken when the *F*2 reached its highest (and where possible, steady) point, excluding transitional values near the boundary which were under the influence of adjacent consonants. All vowels were inspected by hand, and in cases where the Praat formant tracker was inaccurate, hand measurements were made. Tokens were excluded when they were unsuitable for measurement (e.g. the analyst could not see the formants due to poor recording quality or when a phonologically expected vowel was not produced – decided upon both auditorily and acoustically). Formant values in the analyses were in Hertz (Hz).

2.1.4. Descriptive statistics

Descriptive statistics for *F*1 and *F*2 (in Hertz) for the vowels KIT (517 tokens), TRAP (301 tokens), and DRESS (344 tokens) along with information about gender and context are presented in Table 1. The means are also presented visually in Fig. 2.

Table 1
Descriptive statistics (in Hertz).

Vowel	Gender	Context	Mean *F*1	Std. *F*1	Mean *F*2	Std. *F*2
KIT	Female	Neutral	638	95	1805	204
KIT	Male	Neutral	520	63	1579	172
KIT	Female	Australian	625	87	1761	225
KIT	Male	Australian	503	57	1581	170
TRAP	Female	Neutral	647	83	2251	226
TRAP	Male	Neutral	546	66	1877	143
TRAP	Female	Australian	664	99	2244	210
TRAP	Male	Australian	555	91	1883	139
DRESS	Female	Neutral	534	84	2355	233
DRESS	Male	Neutral	473	61	1931	221
DRESS	Female	Australian	542	94	2394	232
DRESS	Male	Australian	468	73	1976	205

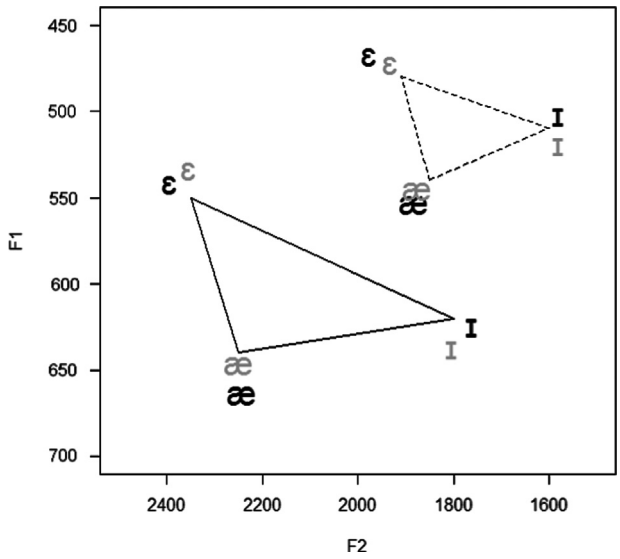

Fig. 2. Mean values for the first and second formants (in Hertz) of vowels in corpus study. Mean values are shown for KIT, TRAP and DRESS in Australian (black) and neutral (gray) contexts. Values for female speakers are connected by the solid triangle. Values for male speakers are connected by the dashed triangle.

Inspection of Fig. 2 shows small differences in the means for all vowels. There are three (non-statistical) patterns that emerge from the means that can be seen for both the males and the females. For KIT, both males and females produce higher (more Australian) variants in Australian contexts. For TRAP, both males and females produce lower (more Australian) variants in Australian contexts. For DRESS there appears to be an effect on *F*2, with both males and females producing fronter variants in Australian contexts. This is contrary to reported results, where Australian variants would actually be backer. The differences in the means are subtle and do not control for effects of underlying factors, including speaker and word-based variation. We therefore now turn to a thorough statistical analysis of the results.

2.2. Statistical analysis and discussion

All data were analyzed using R (R Development Core Team, 2009) and the R packages lme4 (Bates & Maechler, 2009) and languageR (Baayen, 2009; cf. Baayen, 2008). Linear mixed effects models were fit by hand, using model comparison. The dependent variables were *F*1 and *F*2 values for KIT, TRAP, and DRESS words, across separate analyses (i.e. 6 different models were fit). The random intercepts in the analyses were always Speaker and Word (Baayen, Davidson, & Bates, 2008; Clark, 1973). Random slopes for Gender were included for Word.

The primary fixed effect of interest was conversational context. We fit this with three levels: (1) *Australian context*-near an Australian word, where near is operationally defined as within ten words preceding the target utterance, (2) *Australian*-not near an Australian word, or (3) *Neutral context*. This three-way coding was used because the stretches of speech in which Australia was being discussed contain many other Australia-related lexical items, which are not the particular target of analysis. The three levels of this variable allows us to identify: (a) whether Australian conversational contexts elicit more Australian-like speech realizations and (b) whether such an effect is distinguishable from potential lexical priming effects, in which Australian words affect the production of following vowels. In instances where the Australian levels of conversational context were not different from each other, but different from the neutral context, the levels of Australia were combined.

Previous and following phonological contexts were also tested by first including all phonemes as separate levels in factors. Based on the observed groupings of coefficients, we then settled on four binary factors to test: coronal, obstruent, anterior, voiced.

Other fixed effects tested were gender, log word frequency (in the full ONZE corpora – see Hay, Maclagan, & Gordon, 2008),[1] and the number of phonological syllables in the word. Additionally, the following effects were tested and were found non-significant in all models: year of interview and function word identity (whether the critical vowel was from a function word). All non-significant fixed effects were dropped from the model, after first testing for relevant interactions.

2.2.1. KIT: F1 and F2 analyses

Models were fit to 517 tokens of KIT, of which 250 occurred in an Australian context. The final model for *F*1 of KIT is shown in Table 2.

Main effects for gender, following phonological context, syllable length, and conversational context were found. For gender, as would be predicted, males produced lower formant values than females. In addition, a main effect of following phonological context was found for items followed by obstruents, which were lower in *F*1 than when not followed by an obstruent. Also, a main effect of number of syllables in the word was also found; vowels in longer words had lower values of *F*1.

[1] The full corpora were used in order to obtain frequency counts that were local enough to be relevant, yet based on a big enough sample to be relatively robust.

Table 2
Corpus data: fixed effects for model of *F*1 KIT.

	Estimate	Std. error	*t*-Value	*p*-Value
(Intercept)	672.95	19.26	34.94	<0.001
FollowContext Obstruent = Yes	−30.23	7.93	−3.81	<0.001
Gender = Male	−10.41	4.37	−2.38	0.018
Syllable	−100.64	22.68	−4.44	<0.001
Context = Australia	−12.62	5.16	−2.45	0.014

Table 3
Corpus data: fixed effects for model of *F*1 TRAP.

	Estimate	Std. error	*t*-Value	*p*-Value
(Intercept)	640.93	19.58	32.74	<0.001
Gender = Male	−88.11	28.08	−3.14	0.006
ContextAustralia-NearAUS Word = No	2.72	8.05	0.34	0.736
ContextAustralia-NearAUS Word = Yes	26.89	13.57	1.98	0.048

Most importantly, a main effect of the target variable – conversational context – was observed. No significant difference was found between Australian contexts near versus not near an Australian lexical item, and so these levels were combined. Both Australian contexts were significantly different from the neutral context.

When in an Australian context, KIT vowels were approximately 12 Hz lower on *F*1 than KIT vowels in neutral contexts. This small effect is consistent with the hypothesis that people discussing Australia in natural conversations may shift their speech, resulting in more Australian-like realizations. The effect relates to the overall topic of conversation and is not driven by any effect due to particular words given that (1) Australian-associated items were excluded from the analysis and (2) 'word' was included as a random intercept in the analysis.

With respect to KIT *F*2, main effects were found for gender and phonological contexts. For gender, males were found to have lower formant values than females. This is expected, as males tend to have longer vocal tracts. In addition, main effects were found for both previous (obstruent and anterior) and following (coronal and anterior) phonological contexts. Previous obstruent and following coronal contexts were found to have higher *F*2's when present, while both anterior phonological contexts were found to have lower *F*2's when present. No fixed effects of theoretical interest were significant.

2.2.2. TRAP: F1 and F2 analyses

Models were fit to 301 tokens of TRAP, of which 148 occurred in an Australian context. The final model for *F*1 of TRAP is shown in Table 3. Main effects were observed for gender and conversational context. For gender, males were found to have lower formant values than females.

Significantly, a main effect of conversational context was found. When in an Australian context and near an Australian word, TRAP vowels were approximately 26 Hz lower on *F*1 than TRAP vowels in neutral contexts. However, TRAP vowels in an Australian context that were not also near an Australian word fell in between; they were not found to be different from the Australian context near an Australian word or different from the neutral context. The overall result is in line with our prediction, but raises the possibility that additional cues relating to Australia may be necessary to elicit changes in the TRAP vowel.

For TRAP *F*2, main effects were found for gender, phonetic contexts, and syllable length. For gender, males were found to have lower formant values than females. In addition, main effects were found for both previous (voiced) and following (voiced and coronal) phonetic contexts. Previous voiced contexts were found to have lower *F*2's when present, while both following contexts were found to have higher *F*2's when present. Also, a main effect of number of syllables in the word was also found, in which vowels in longer words had lower values of *F*2's. No fixed effects of theoretical interest were significant.

2.2.3. DRESS: F1 and F2 analyses

Models were fit to 344 tokens of DRESS, of which 169 occurred in the Australian context. Main effects for DRESS *F*1 were observed for phonological context, syllable length, and log lexical frequency. For phonological contexts, obstruents that were present prior and following the critical vowel were found to lead to lower values for *F*1. Also, words with longer syllable lengths were found to have higher values of *F*1. Finally, for log lexical frequency, frequent words had higher *F*1. This is consistent with work on the ongoing raising of DRESS, which shows that frequent words lag in this change (Hay, Pierrehumbert, Walker, & LaShell, under review). Gender and conversational context were not significant.

The DRESS *F*2 analysis revealed significant main effects for gender, phonological context, and conversational context. A main effect of gender was found, where males produced lower formant values than females. The following phonological context, when it was anterior, was found to lead to higher values of *F*2. Of primary interest, an effect of the conversation context was observed. No difference was found within the Australian contexts, regarding lexical priming, but both Australian contexts were significantly different from neutral, and so were collapsed together. The Australian context predicted higher values of *F*2 than the neutral context. This

Table 4
Corpus data: fixed effects for model of *F*2 of DRESS.

	Estimate	Std. error	*t*-Value	*p*-Value
(Intercept)	2256.42	61.78	36.52	<0.001
FollowContextAnterior = Yes	99.55	32.95	3.02	0.003
Gender = Male	−403.03	80.44	−5.01	<0.001
Context = Australia	51.92	18.24	2.85	0.005

suggests that people discussing Australia in natural conversations are shifting their speech, resulting in *less* Australian-like DRESS vowels, compared to normal, which is at odds with our prediction. The Australian context is related to a shift of approximately 52 Hz. The final model for DRESS *F*2 is shown in Table 4.

2.3. Measurement verification

As outlined above, the corpus analysis revealed some small, but intriguing, effects of whether the conversational topic related to Australia. The small nature of the effects caused us to question the degree to which they were reliable, as opposed to being small artefacts relating to measurement error or bias. The original measurements were taken manually by the third author, who was not blind to the hypothesis, and who was the same analyst who identified the contexts as 'Australian' or 'not Australian'. Thus, she was always aware of the condition when taking the formant reading. In order to safeguard against potential error or bias, we thus removed all coding relating to topic of conversation, and provided each token to "blind" research assistants, to provide manual formant readings following the same procedure as the original analyst. All tokens were reanalysed, and three different analysts were involved in this task. The correlation between the original readings and the new readings was significant and high (Spearman's rho = .85 for overall *F*1, .88 for overall *F*2; when analyzed separately by each vowel/formant, values span .82–.87, all correlations highly significant). We created conservative subsets of the data, comprising only tokens for which the two measurements were within 50 Hz of each other, for *F*1, or 100 Hz of each other for *F*2. This process removed all outliers and those tokens for which readings were apparently not clear. We then repeated our analysis on the new measurements in these data-sets. The significant effects relating to contextual mention of Australia reported in Section 2.2 all remained statistically significant in this analysis.

2.4. Corpus analysis: discussion

We hypothesized that the short front vowels KIT, TRAP, and DRESS would have a slightly more Australian realization when produced in an Australian context. This idea was supported by small significant effects for KIT and TRAP on *F*1. Variation in KIT exemplifies the most salient difference between New Zealand and Australian English, and is accurately identified by speakers of both NZE and AE. This finding is in line with speech perception research that has found KIT to be susceptible to priming shifts (Hay et al., 2006; Hay & Drager, 2010). It is perhaps not surprising, then, that the effect of the conversational context is observable with this vowel in natural speech.

In addition, we also observed speech changes for TRAP in the direction of an Australian realization when in an Australian context. Interestingly, this effect was only found when the utterance was near an Australian related term. This result may be related to the lesser saliency of TRAP in the consciousness of New Zealanders, as reported by Ludwig (2007). As the differences in TRAP are not as readily accessed as KIT, the presence of a nearby Australian term may serve as an additional prime to Australia, facilitating Australian-like speech realizations for this (comparatively) lesser known vowel difference.

Our hypothesis was not supported by the results found for DRESS. Australian realizations of DRESS are fronter (i.e. have a higher *F*2) than New Zealand realizations (e.g. Cox & Palethorpe, 2008; Maclagan & Hay, 2007), though the extent of this difference has varied over the last few decades, as vowels on both sides of the Tasman have been on the move (Cox & Palethorpe, 2008; Maclagan & Hay, 2007). Curiously, we found a result which goes in the opposite direction. DRESS realizations were significantly more 'Australian' in the *neutral* context compared to the Australian context.

How can we explain this apparently contradictory result? We present two possible explanations, which we test in a later experiment.

(1) False stereotypes: One clear difference between KIT and DRESS is that variation in KIT is reliably above the level of consciousness for New Zealanders, but this is not the case for DRESS. As mentioned, Ludwig (2007) found that NZE speakers are accurate when identifying AE KIT as AE, but incorrectly, and importantly in light of our results, tend to identify AE DRESS as NZE. Ludwig (2007:104) speculates that this misattribution may relate to the social variation of DRESS in New Zealand. "DRESS is moving to a high front position to the extent that in some cases it overlaps with FLEECE (McKenzie, 2005) and it is possible that its higher variants are becoming stigmatized. It is possible that New Zealanders perceive the lower variants as New Zealand-like because they have also experienced older New Zealanders producing them".
Whatever the reason, it appears that some NZers misattribute their high front DRESS as belonging to Australians. Thus, one possible mechanism for the contradictory corpus result is that it is due to false stereotypes. Individuals adopt the vowel realization they associate more with Australia, when in the Australian context.

(2) Cross vowel-priming: As discussed, the high front KIT vowel is a stereotype of Australian English, while differences in the realization of DRESS are less salient. Given that the dialectical difference of KIT is so salient, it may be the case that people generalize the high front location of KIT to other vowels, leading to a higher and fronter articulation for these other vowels. In this sense the high front AE KIT region may attract nearby vowels, such as DRESS, when discussing Australia. By this hypothesis, the observed effect would not be driven by any overt false stereotypes or beliefs about DRESS, but would simply be a cross-vowel priming effect that occurs in real time, as the high front part of the vowel space becomes primed by the mention of Australia. Evidence in support of this idea can be found in Tilsen's (2009) cross-phoneme priming study, which investigated the extent to which the qualities of a perceived vowel could affect the production of other vowels. Tilsen found that the recent presentation of a centralized vowel (e.g. modified /a/ and /i/) lead to more centralized productions of /a/ and /i/, crediting the recent exposure of the centralized stimuli as attracting and informing the subsequent productions. Thus, in relation to the current study, note that the actual New Zealand production of DRESS is much closer to the Australian KIT vowel than the New Zealand production of KIT. Thus, it is possible that DRESS realizations may be affected by such a mechanism, where DRESS productions are affected by KIT through cross-phonemic vowel priming.

These explanations are, of course, speculative and represent post-hoc attempts to explain a result patterning in the opposite direction from what we predicted. Additionally, our observed result may also be due to random chance and might not be replicated in another sample. We do not have another available spontaneous sample in which to test these explanations. Instead, we turn to an experimental paradigm in order to test our DRESS hypotheses.

It should also be noted that the results from the corpus analysis are significant, but small. The small size is of theoretical interest, since it would seem to support the interpretation that these are automatic, subtle, exposure-based effects, and are not related to overt style-shifting or identity display – both of which we might expect to lead to larger effects. This subtle effect is in line with the size of the effects found in accommodation research (e.g. Babel, 2010, 2012; Nielsen, 2011).

In casual speech there are many things happening simultaneously, not all of which can be controlled. We have done our best to take care of as much variation as possible statistically, to see if our small effects still remain. However a more controlled follow-up would be reassuring. When the lexical items and context is tightly controlled, can an 'Australian' context cause adjustments in New Zealanders' speech?

3. Australia priming experiment

The primary aim of this experiment was to replicate and clarify the corpus DRESS results by: (1) Assessing the realization of DRESS words in Australian and non-Australian primed contexts; would the unexpected effect of an Australian context on DRESS be replicated? (2) Investigating a possible cross-vowel priming effect due to the presence of KIT words; does recent exposure to an Australian-primed KIT vowel increase the observed effect on DRESS? (3) Identifying whether a speaker's experience and/or attitudes are also factors in the speech production process. If the unexpected result with respect to DRESS is mediated by speaker stereotypes, then individual differences in exposure to and/or attitude to Australian English may play a role.

A simple word-reading experiment was conducted. Participants read words one at a time from a computer screen. Target words were DRESS words found in various conditions, involving the manipulation of the two preceding words. The immediately preceding word (referred to as position 2) could be a KIT word, an Australian-related word, or a non-KIT-non-Australian (neutral) word. The word before that (position 1) could be either Australian or non-Australian. The full design is outlined below.

3.1. Method

3.1.1. Participants

Thirty-one NZE speakers (19 females, 12 males) participated in the experiment. Participants were recruited from the University of Canterbury via posted fliers on campus and were compensated with a $10 voucher. Participant ages ranged from 18 to 39 years (Mean = 23.3) and reported no history of speech or hearing impairments. Students were not screened for their major specialization or classes attended.

3.1.2. Stimulus materials

Seven different word sets were created to fulfill the experimental requirements for the full word list. All words and frequencies (except those on the Australian lists) were obtained from CELEX. Words relevant to this experiment can be seen in Appendix B.

3.1.2.1. Australian words. There were two Australian word lists, Australian words for position one (AUS1) and Australian words for position two (AUS2). There were 18 AUS1 words and six AUS2 words. Both sets of words were considered to be closely associated with Australia. In terms of associative strength, however, words in AUS1 were considered more closely associated, though the intention of both lists was to be very highly associated with Australia to suitably prime the idea of Australia. Words in AUS1 contained names of states, major cities, and regions, the country name itself, names referring to the country's peoples, stereotypical animals, and the name of the country's rugby team (also a stereotypical animal) in Australia. Words in AUS2 contained the name of a popular beach, less-stereotypical animals than those in AUS1, a country- unique object, and a slang term for the country. AUS1 words were

always in the first position of a critical Australian trial, whereas AUS2 words were presented in the second position of a type of critical Australian trial.

3.1.2.2. KIT words (KIT). KIT words consisted of 15 monosyllabic words with the KIT vowel with a word frequency less than 50 per one million.

3.1.2.3. DRESS words (DRESS). DRESS words consisted of 36 monosyllabic words with the DRESS vowel with a word frequency less than 100 per one million.

Low frequency words were selected to test our predictions. In the accommodation literature, low frequency words are typically used to test a variety of variables in phonetic accommodation as they have been found to be more susceptible to shifts in pronunciation of previously heard (trained) words (e.g. Goldinger, 1998; Goldinger & Azuma, 2004; Miller et al., 2010; Sanchez, 2011) and can also generalize to previously unheard (untrained) words (Nielsen, 2011), which is particularly relevant to our current study.

3.1.2.4. Neutral words. There were two neutral word lists, neutral words for position one (Neutral1) and neutral words for position two (Neutral2). There were 18 Neutral1 words and 15 Neutral2 words. Neutral1 words were created as pairs or neutral equivalents to the AUS1 words. Neutral2 words were monosyllabic words composed of various vowel types (excluding KIT, TRAP, and DRESS).

3.1.2.5. Filler words. 324 filler words (varying in word frequency, not controlled) were composed of words varying in length (from one to four syllables) that did not refer to Australia or contain KIT, TRAP, or DRESS vowels. There were three times as many filler words than all other words (listed above) combined (324 filler to 108 all other words; 432 words in total). The rationale for this difference in proportion was to place a certain amount of distance between any critical sequence of words. As such, nine randomly selected filler words were presented before each critical sequence (explained further in Section 3.1.3).

Preexisting questionnaires that measured both the participant's experience with and explicit attitude towards Australia were completed at the end of the experiment (see Drager, Hay, & Walker, 2010). The questionnaires are given in Appendices C and D. The experience measure assessed factors such as, whether they have visited Australia, length of stay, whether they watch Australian television, and whether they socialize with Australians. The explicit attitude measure collected agreement ratings with statements relating to desires to holiday in Australia, the degree to which New Zealanders and Australians are similar, Australia's human rights records, and the degree to which stereotypes relating to Australia are true.

For the experience section of the questionnaire, each of the six experience questions was scored between 1 and 4, meaning that the summed scores could range from 6 to 24. The summed scores were converted into a number between 0 and 1 using the following formula: ((Participant summed score) −6)/18). This anchored the potential scores at 0 (for the lowest possible score) and 1 (for the highest). The median score was .33.

The highest score was .667. This participant has been to Australia five times, with each visit lasting approximately two weeks. She has an Australian Aunt and several Australian friends, and is in contact with them on an average of three times per week. She watches two specific Australian television shows regularly and watches television in general, ‘sometimes’. She also follows international sports, specifically netball and rugby.

There were five participants (three females, two males) who shared the same lowest experience score (.056). These participants never visited Australia and do not know any Australians. These participants did not list any specific Australian television shows they watch regularly. Some watch television more generally, and some follow international sports.

For the attitude section of the questionnaire, a similar process was followed. Participants were scored on their responses to the statements, and the summed scores were converted to an index that could span from 0 to 1. The actual scores varied from .267 to .767, with a median of .467.

3.1.3. Procedure

The experiment was conducted in a small quiet room dedicated to running experiments. Participants were run in the experiment, one at a time, by a single experimenter (the first author, who is a speaker of American (Californian) English). The production experiment lasted approximately 25 min and a total of 432 words were said aloud and recorded. Subsequently, an Australian Questionnaire was administered. All participants were debriefed on the purpose of the experiment before leaving.

The experiment was presented using E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, Schneider, Eschman, & Zuccolotto, 2002a, 2002b). The first half of the experiment consisted of filler words and critical baseline sequences, while the latter half of the experiment consisted of filler words and critical experimental sequences. All baseline sequences preceded experimental sequences, in order to avoid any potential longer-term priming of the Australian Context. All critical sequences were preceded by nine randomly selected filler words without replacement. Words in the critical trials were randomly selected from the appropriate word lists without replacement. A given critical sequence was randomly selected from a list of possible sequences and was reset after all possible sequences were used. There were two critical baseline sequences: Neutral1-KIT-DRESS and Neutral1-Neutral2-DRESS. There were three critical experimental sequences: AUS1-AUS2-DRESS, AUS1-KIT-DRESS, and AUS1-Neutral2-DRESS. All words were presented one at a time with a 2.5 s inter-stimulus interval. Example sequences are given in Table 5.

Participants were fitted with a head mounted microphone (Beyerdynamic Opus 55 Mk II). Each word was recorded directly onto the computer (via a Sound Devices USB Pre 2 preamp). Participants were told that their task in this experiment was to say the words they would see on the computer screen “out-loud, quickly, but clearly”. They were told that at the mid-way point of the experiment that

Table 5
Examples of baseline and experimental sequences. Baseline sequences all occurred before experimental sequences. Nine randomly selected filler words intervened between each critical baseline or experimental sequence.

Position 1	Position 2	Position 3
Critical baseline sequences		
Neutral1	KIT	DRESS
Marmalade	*Skit*	*Peck*
Neutral1	Neutral2	DRESS
Path	*Sleep*	*Nets*
Critical experimental sequences		
AUS1	KIT	DRESS
Adelaide	*Pip*	*Mesh*
AUS1	Neutral2	DRESS
Perth	*Game*	*Sled*
AUS1	AUS 2	DRESS
Koala	*Platypus*	*Vest*

Table 6
Experiment data: descriptive statistics (in Hertz).

Vowel	Gender	Context	Mean *F*1	Std. *F*1	Mean *F*2	Std. *F*2
KIT	Female	Neutral	570	101	1915	184
KIT	Male	Neutral	485	62	1624	156
KIT	Female	Australian	580	99	1897	192
KIT	Male	Australian	476	59	1611	172
DRESS	Female	Neutral	441	55	2531	255
DRESS	Male	Neutral	416	56	1987	190
DRESS	Female	Australian	447	61	2531	248
DRESS	Male	Australian	410	63	2038	210

the computer would instruct them to take a break and have some water before continuing with the rest of the experiment, which they controlled by pressing the space bar.

3.1.4. Acoustic measurements

The first and second formants (*F*1 and *F*2, respectively) of the KIT and DRESS words were measured by hand using Praat. The measurements were taken when *F*2 was at its peak, excluding transitional values near the boundary which were under the influence of adjacent consonants. In cases where the Praat formant tracker was inaccurate, hand corrections were made. Words that were mispronounced were not included in the data (i.e. two KIT tokens and 29 DRESS tokens in total). In addition, the word “weld” was removed from all analyses, due to the fact that the vowel occurred pre-laterally (see Bauer, 1986 for the influence of /l/ on preceding vowels in New Zealand English).

3.1.5. Descriptive statistics

Broad descriptive statistics for *F*1 and *F*2 (in Hertz) for the vowels KIT (378 tokens), and DRESS (999 tokens) along with information about gender and context are presented in Table 6. No particular trends are discernable from the means. We therefore turn to a statistical analysis of the patterns.

3.2. Results and discussion

As with the corpus analysis, data were analyzed with linear mixed effects models. The random intercepts in the analyses were Speaker and Word. Word also had a random slope for gender (Baayen et al., 2008; Clark, 1973). The predicted variable consisted of *F*1 and *F*2 values DRESS words. The fixed effects considered consisted of Speaker, Gender, Phonological Context, Position 1 (Australian or Neutral), and Position 2 (Australian, KIT, or Neutral), as well as aggregate scores on Australian Attitude, and Australian Experience scales. No correlation was found (correlation $=-0.074$) between the Australian Attitude and Australian Experience scores obtained from the post-questionnaire and were thus included as separate effects. Non-significant fixed effects were dropped from the model, after testing for possible interactions. The position 2 KIT words were also analyzed in the same way, with random intercepts for Speaker and Word, and fixed effects of Speaker, Gender, and position 1 (Australian or Neutral).

Table 7
Experimental data: fixed effects for model of *F*1 DRESS.

	Estimate	Std. error	t-Value	*p*-Value
(Intercept)	470.84	13.24	35.58	<0.001
FollowContextVoiced = Yes	−27.98	7.91	−3.54	0.001
FollowContextCoronal = Yes	−14.39	6.36	−2.26	0.030
FollowContextObstruent = Yes	−20.11	8.9	−2.26	0.030
Position2isAus = Yes	2.82	3.24	0.87	0.384
AusExperienceRank	14.93	6.84	2.18	<0.001
Position2isAus = Yes: AusExperienceRank	11.6	3.36	3.46	<0.001

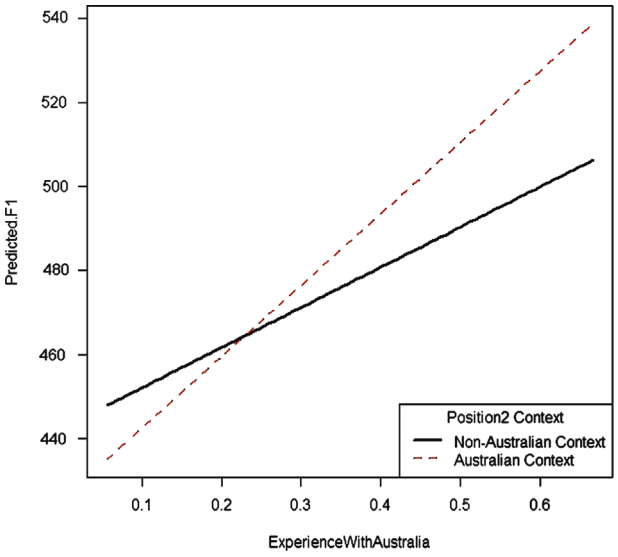

Fig. 3. The interaction between Experience Score, and whether position 2 is Australian. Speakers with greater experience with Australia show a shift in the predicted direction. Speakers with less experience with Australia do not.

3.2.1. DRESS: F1 and F2 analyses

Models were fit to 999 observations of DRESS. Like the corpus data, Gender was not significant for DRESS *F*1. For *F*1, main effects were observed for phonological context. The following phonological context when voiced, coronal, or obstruent were found to have lower *F*1 values. Of particular interest, a significant interaction was observed between Position 2 and Australian Experience for *F*1 (see Table 7, Fig. 2). The analysis revealed that the KIT primed items were not significantly different from the neutral items. In the final model we combine KIT and neutral items to test the overall difference between 'Australian' and 'non-Australian' position 2 contexts. In this analysis, there is a significant interaction between position 2 context and the participant's experience with Australia. It was found that participants with a greater amount of experience with Australia shifted their DRESS realization when an Australian related word preceded it, resulting in a more Australian-like realization.

Note that the design is not completely orthogonal. Position 2 is only an Australian item when position one is also Australian. We therefore conducted further models of subsets of the data, in order to determine the true source of this effect. First, we restricted the dataset to the cases where position 2 is not Australian, allowed position 1 to vary and tested for an interaction with experience. The interaction between position 1 and Experience, when restricted to the dataset where position 2 is not Australian, did not reach significance. In other words, the 'Australia1-NonAustralia2-DRESS' vowels did not differ significantly from the Neutral1-NonAustralia2-DRESS vowels. Second, we restricted the dataset to those cases where position 1 is Australian, and allow position 2 to vary and tested for an interaction with experience. Here, position 2 and Experience significantly interact. In other words, the 'Australia1-Australia2-DRESS' vowels did significantly differ from the 'Australia1-NonAustralia2-DRESS' vowels.

Taken together, this suggests that the important contrast in this dataset is between the Australia1-Australia2-DRESS trials and all the others. The double Australian prime affects speakers' *F*1 DRESS value, in a way that is influenced by their experience with Australia. This result is shown in Fig. 3. An Australian realization for DRESS has higher *F*1 values compared to NZE. The more experience our participants had with Australia, the more likely they were to produce higher *F*1 values for DRESS in the Australian context. However, participants with lowest levels of experience did not shift in this direction. If anything, they trend in the opposite direction, much like the speakers in the corpus. The existence of this interaction explains why we did not see any trends in the raw means. Different speakers are behaving in different ways with respect to this variable, limiting the value of basic overall means.

The *F*2 results found main effects for gender and phonetic context. The gender main effect revealed that male speakers produced lower *F*2 than female speakers. Following phonological context, when it was obstruent, was found to have lower values of *F*2. No fixed effects of theoretical interest were significant.

3.2.2. KIT: F1 and F2 analyses

Models were fit to 378 tokens of KIT. For KIT *F*1 and *F*2, a main effect was found for speaker gender, where males were found to produce lower values than females. KIT *F*2 also revealed a main effect for phonological context, where previous anterior contexts were found to have lower formant values. The experiment therefore does not replicate the result from the Corpus where KIT vowels were produced with an Australian realization when discussing Australia. There are likely two reasons for this result. First, the experiment was primarily designed to investigate the behavior of DRESS, because of this there are fewer tokens of KIT available for analysis than DRESS. Second, however, note that the significant effect for DRESS was driven by the double primed Australia–Australia priming condition. This condition was not tested for KIT, which was only ever preceded by a single Australian item. The experimental context is, of course, different from the corpus context, in which the focus of the discussion is Australia-related, and so the entire local context is somewhat primed. In the experiment we are dealing with lexical, rather than broad contextual priming. Priming by a single preceding item may simply be insufficient to elicit the effect.

3.2.3. Experimental analysis: Discussion

The experiment revealed three results of interest. (1) An overall effect of Australian priming on DRESS vowels was found. (2) The presence or absence of a KIT vowel had no effect on DRESS realizations, discounting the cross-vowel hypothesis as a likely explanation for the direction of DRESS differences. (3) Speakers with a reasonable amount of exposure to Australia shift in a more Australian manner (as predicted), while speakers with little experience show no such effect.

As this investigation makes use of the idea of ‘concept activation’ (e.g. priming), though specifically with words concerning Australia, it should be noted that there is also a possibility that the experimenter (a speaker of American English) may have served to alter the subjects' speech (e.g. Hay, Drager, & Warren, 2009). However, given the pattern of results, it seems unlikely that the experimenter had a lasting effect, if any. Although American speech for KIT and DRESS is more similar to AE speech than NZE speech, the speech produced during the baseline portion of the experiment would have been most affected which would only serve to mitigate the expected result. As it stands, the main result we find is that when *two* Australian words precede a DRESS target word, the DRESS word is pronounced in a more Australian manner, but specifically for those who have considerable experience with Australia.

Recall the differences between our highly experienced and inexperienced subjects' knowledge relating to Australia. Those with much Australian experience were exposed to multiple sources of Australian speech while those with little Australian experience were exposed to few, if any, sources of Australian speech. Finding that speakers' experience with Australia can affect DRESS vowel realizations when Australian-related words precede it perhaps sheds some insight on the unexpected direction of the corpus results. The experiment sought to differentiate between the possibility that the corpus results were obtained either by cross-vowel priming or stereotypes. The experimental results do not provide evidence for cross-vowel priming. However, the stereotype hypothesis may still be viable.

If the corpus results were due to NZE speakers with the false stereotype that Australians produce higher/fronter DRESS vowels than New Zealanders, then we should expect to see a similar effect to that shown in Fig. 2. We would expect that in an Australian context, speakers with more experience of AE may be more likely to shift toward actual Australian realizations; whereas speakers with little experience may shift in the direction they (falsely) perceive to be more Australian. We therefore revisit the corpus data to test this possibility.

4. Canterbury Corpus, revisited

In light of the experimental results, the corpus was revisited to identify if additional factors, such as experience with Australia, played a role in the speech productions.

4.1. Method

4.1.1. Participants and materials

The Canterbury Corpus data used in the first study (Section 2) was used.

4.1.2. Procedure

Additional notes on the corpus transcripts were made concerning the speakers' attitude and perceived experience with Australia. Attitude toward Australia was not used in the subsequent analyses as the variation of attitudes was not deemed to be informative; all speakers were considered neutral toward Australia, except two, where one had a positive attitude, while the other was considered to have a negative attitude toward Australia.

However, an experience measure was included in the analyses. The 15 speakers were given a score on a scale from 1 to 100, where 1 = least Australian experience and 100 = most Australian experience. Placement on the scale reflects the overall impression of experience with Australia, given the information provided in the transcripts. Items considered in this judgement were things such as the amount of time the speaker had spent in Australia, whether the speaker personally knew any Australians, whether they regularly watch Australian television shows, and the number of different places mentioned in Australia, to keep in line with the questions answered by the participants in the experiment. A scale of 100 was used in order to allow a full ranking of participants to emerge (which would not be possible, for example, with a scale of 10), and for the scale to contain enough freedom for the distance to vary

considerably between adjacent participants on the scale, if appropriate. The placement on the scale was subjective, and was completed by an independent research assistant on the basis of the full transcripts of the speakers' interviews. The research assistant did not listen to the audio, and was blind to the purpose of the task. The scores were reviewed and validated by the first author.

To illustrate the differences between the experiences of the speakers we provide the following. The speaker rated as having the most experience with Australia was in her mid 40s at the time of the interview, and described having lived and worked in Australia for several years during her 20s. The speaker with the median score reports that he had only ever been overseas once, for a three week work trip to Australia. The speaker who was rated as having the least experience did not report having visited Australia during the course of her interview. In her 'Australian' section she reports that she likes to watch Australian television programmes, and discusses the details of a number of these shows. The full set of scores assigned, from highest to lowest, was 100, 96, 92, 88, 84, 70, 66, 58, 50, 42, 34, 26, 20, 8, 4.

This ranking is, of course, not precise, as it was made based on partial and complete information – just that information that speakers happened to impart during their interview. It is, however, the best that can be done with the information we have available. We took this approximate ranking and tested whether it significantly improved the models described in Section 2 – either independently, or in interaction with conversational context.

4.1.3. Acoustic measurements

The measurements used in the first corpus study (see Section 2.1.3) were used. No new measurements were conducted.

4.2. Results and discussion

We added the speakers' Australian experience score into each of the models reported in Section 2, for which we reported a significant context effect. We then tested whether these significantly improved the models. The main effect of Australian experience was tested, together with interactions between Australian experience and conversational context. The speakers' Australian experience did not significantly improve the models for KIT or TRAP. However, the inclusion of Australian experience did significantly improve the models of both *F*1 and *F*2 for DRESS.

DRESS *F*1 did not originally show any effect of conversational context and the addition of the experience rank measure did not change this, but our new variable, experience, was found to have a significant effect. The result is the model shown in Table 8.

Interestingly, there is a significant main effect of Australian experience on the overall value of *F*1 for DRESS. Speakers with more experience of Australian English have higher values of *F*1, or – in other words – have naturally more Australian realizations of their DRESS vowels. This may be chance (keeping in mind that this is a cross-speaker comparison, and we are dealing with 15 speakers), or it may indicate some subtle long-term effect of exposure to Australian English in those who have had the most exposure. For example, as discussed, the top ranked speaker had spent some time in her 20s living in Australia. Although she was in her 40s when being recorded, it is possible that this time had some residual effect on her vowels.

For DRESS *F*2, an interaction was found between Australian experience and conversational topic. This model is shown in Table 9. The interaction between conversational context and Australian experience is shown in Fig. 4. We repeated this model using an independently acquired set of measurements, as described in Section 2.3. The interaction between experience and context remains significant.

Table 8
Corpus data: fixed effects for model of *F*1 DRESS.

	Estimate	Std. error	*t*-Value	*p*-Value
(Intercept)	432.73	28.03	15.44	<0.001
PreviousContextObstruent = Yes	−26.95	7.62	−3.54	<0.001
FollowContextObstruent = Yes	−25.57	7.75	−3.30	0.002
LogFrequency	4.72	1.82	2.59	0.012
Syllable	8.98	4.62	1.95	0.054
AusExperienceScale	0.9	0.33	2.73	0.015

Table 9
Corpus data: fixed effects for model of *F*2 DRESS.

	Estimate	Std. error	*t*-Value	*p*-Value
(Intercept)	2285.27	92.68	24.66	<0.001
FollowContextAnterior = Yes	85.76	32.69	2.62	0.010
Gender = Male	−418.07	76.22	−5.49	<0.001
Context = Australia	202.52	41.81	4.84	<0.001
AusExperienceScale	−0.27	1.22	−0.22	0.831
ContextAustralia:AusExperienceScale	−2.48	0.62	−3.98	<0.001

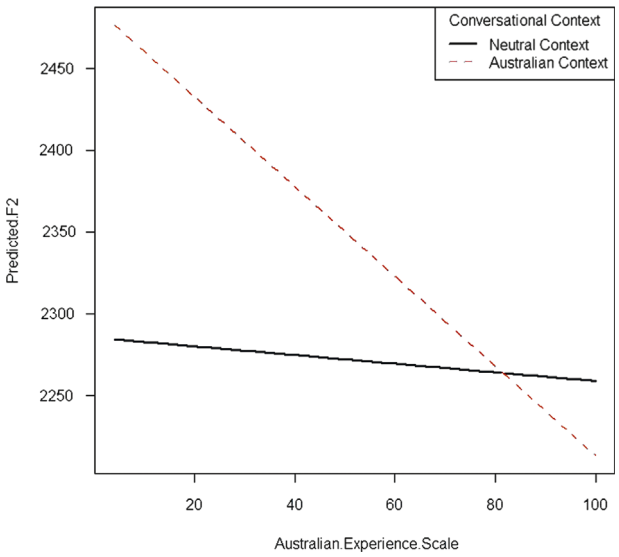

Fig. 4. The interaction between speakers' experience with Australia (least experienced = 1, most experienced = 100) and the predicted *F*2 for DRESS. Separate lines are shown for the Australian (dashed) and non-Australian (solid) contexts.

Clearly, the score given by the independent research assistant for experience with Australia is not predictive of the *F*2 of the DRESS vowel in non-Australian contexts. However, there is a reliable association between *F*2 and the speaker's experience with Australia when discussing Australia. Speakers with the least experience of Australian English show the greatest degree of shift in *F*2 when in the Australian context. This shift, as found in the original corpus analysis, results in *less-Australian* like realizations. However, given their relative lack of experience with Australia, it is possible that a false stereotype of a high front DRESS (inappropriately) belonging to Australians was activated. Speakers with the most experience of Australian English do not display this large shift and, if anything, go in the opposite (originally predicted – actually Australian-like) direction. Note that if we replace the experience score with a ranking, the same overall result is found.

This result is in line with the experimental findings. Those who have high levels of experience with Australia produced more Australian speech when in an Australian context compared to those with low levels of experience with Australia. This result also puts the original corpus study in a new light. The inclusion of Australian experience was able to uncover a difference between Australian and Neutral contexts in the predicted direction. Those with more Australian experiences likely have more stored exemplars of actual Australian realizations of speech, including the DRESS vowel, which may have facilitated the realization of true Australian-like DRESS vowels in Australian contexts, as opposed to activating a false stereotype of AE DRESS.

5. Discussion and conclusion

This paper investigated whether the concept of Australia can affect phonetic realizations of the short front vowels KIT, TRAP, and DRESS. Both the corpus and the experimental work together suggest that the answer is ‘yes’. Small but significant effects emerge across four linear mixed effect models, and two disparate sources of data. The effects are not large, but nor should we expect them to be. They are not identity related, but simply emerging as a subtle automatic consequence of the socioindexical labeling of selected exemplars.

The combined picture provides evidence of patterns of contextual influence on vowel production. But, if one were to look at the two studies independently, one's interpretation of exactly which vowels are affected, and which *aspects* of these vowels, including the direction of change would differ considerably.

In the corpus study, we found a small non-interacting effect of the conversational topic (Australian or non-Australian) on KIT *F*1 and TRAP *F*1, in the expected direction, where the Australia-related talk was associated with more Australian-like *F*1 realizations. This did not interact with individuals' experience with Australia.

Given its salience, KIT is certainly the vowel one would expect to observe the most robust production result, as it represents the most widely known linguistic difference between New Zealand and Australian English, and has previously been shown to be reliably malleable in speech perception tasks (c.f. Drager & Hay, 2010; Hay et al., 2006). Interestingly, Drager et al. (2010) also found an effect of Australia on KIT word list productions, but their effect was present on *F*2, rather than on *F*1 (as in the corpus result).

TRAP, though less salient than KIT, is accurately identified with respect to the dialects and was found to shift according to conversational context in the appropriate direction. However, this effect was only observed when additional cues were present, such as when an Australian word was near. This suggests that for lesser known dialectical differences, additional primes relating to the dialect are necessary to observe shifts in speech toward the primed dialect.

DRESS, on the other hand, is not only less salient, it is inaccurately identified with respect to the dialects and was found to shift according to the conversational context in an unexpected direction. Here, speech about Australia was associated with lower *F*2's for DRESS – i.e. *less Australian-like* variants. Puzzled by this result, we set out to try and investigate it experimentally.

In an experiment devoid of natural conversational context, would Australian-associated lexical primes affect vowel production? The answer that our experiment gives is 'yes', but that a single recent Australian prime is not enough, which bears similarity to our corpus TRAP results. Our experiment revealed that *F*1's for DRESS vowels followed by a double lexical Australian prime (i.e. two Australian-related words in a row) were realized differently from other DRESS vowels. In addition, the lexical context preceding DRESS also interacted with the speakers' experience. For speakers with moderate to high exposure to Australian English, the effect went in the originally hypothesized direction, meaning that Australian-primed DRESS vowels had more Australian-like productions. However this effect was absent for those with less exposure to Australian English, and even trended the opposite way.

Finding an effect of experience with Australia on DRESS realizations in the experiment motivated us to revisit the corpus results. When adding a posthoc measure of the corpus speaker's experience with Australia to the analyses, a complementary effect on *F*2 DRESS was revealed. Speakers who were less experienced with Australian English were found to produce less Australian-like vowels in the Australian context, whereas speakers with more experience trend in the opposite (predicted Australian-like) way.

The experimental and corpus results for DRESS are thus aligned; speakers with a high level of Australian experience produce more Australian-like DRESS vowels in Australian-primed contexts, while speakers with less Australian experiences produce less Australian-like DRESS vowels in Australian-primed contexts. Nonetheless, the experimental results and the corpus results for DRESS also differ, in that (1) the corpus result is on *F*2 DRESS, and experimental result is on *F*1 DRESS, and (2) for the experiment, the largest effect of context is observed with the most experienced speakers, while for the corpus, the largest effect of context is observed with the least experienced speakers. The differences obtained from the experiment and corpus are discussed below.

With respect to the first point of difference between the experiment and corpus results, we do not have an explanation as to why the DRESS corpus result reaches significance on *F*2 while the experimental results reach significance on *F*1. We note, too, that our corpus KIT result was found on *F*1, whereas previous work with wordlists report an effect on *F*2 (Drager et al., 2010). Although differences in *F*1 and *F*2 were found in different contexts for both DRESS and KIT it should be highlighted that both KIT and DRESS differ in *F*1 and *F*2 for both NZE and AE. The studies herein differ in a number of dimensions which could cause one or other of these to be more salient. Specifically, when comparing the corpus DRESS and the experimental analysis of DRESS, there is an obvious task-based difference that may affect speech rate and clarity. In addition, word frequency may also play a role, as experimental stimuli were restricted to low frequency words, whereas the corpus analysis was not, and indeed was likely to contain an overrepresentation of high frequency words (by virtue of their frequency). Furthermore, the corpus data was collected 10 years earlier than the experimental data. This is an important difference, as the DRESS vowel changed in both dialects, becoming more distinct over this time – particularly in *F*1 (Cox & Palethorpe, 2008; Maclagan & Hay, 2007). Thus, there may be task related issues leading to the effect on different formants using different methods, but this may also relate to the increased difference between NZE and AE at the time of the recordings, especially on the *F*1 dimension. What is clear is that one should hold back from drawing conclusions about the larger relevance of *F*1 vs. *F*2, based on the results of any single study using any particular methodology.

We note, too, that the previous perceptual work on NZE short front vowels has not found priming effects with DRESS, but has reliably found effects with KIT and found some results with TRAP. This may be simply due to the experimental designs used in the previous studies, which collected more data for KIT (the primary object of investigation) than DRESS. Alternatively, reliable priming effect on DRESS might not be found if analyses do not consider the participant's notion of dialectical DRESS differences, as false stereotypes may be activated. In a similar vein, it may be that speaker experience with Australian English needs to be incorporated into the DRESS priming analyses to allow an effect in perception to emerge, as it emerged here with production.

With respect to speech production, the current DRESS result is in line with Babel's (2010) phonetic convergence findings. She found that New Zealanders shadowing an Australian voice were more apt to shift towards an AE DRESS realization than for the other vowels tested. Babel (2010:453) suggests that the lack of salience of this variable in New Zealand English may have facilitated the degree of shifting in this context, where (unlike in our own studies) an explicit phonetic model is provided, arguing that "…with a large amount of phonetic space between the NZE and AE categories, there is ample room for a New Zealander to accommodate, and the lack of salience for this category may facilitate the accommodation process". Notably, while Babel does show some effect of speakers' implicit attitudes on the degree of shift of some vowels, the implicit attitudes do not seem to influence her participants' DRESS productions. There is clearly and important interplay between the nature of the task, speaker experiences and/or attitudes, and the stereotypes and salience of particular variables. Only by conducting studies across a range of methodologies and variables will we be in a position to properly disentangle these factors.

With respect to the second point of difference between the experiment and corpus results, we are left asking why the DRESS effects are carried by the speakers with most Australian experience for the experiment, and by the speakers with least Australian experience for the corpus. Again, this cannot be answered definitively. This difference likely relates to the different speaker populations and recording times of the two data-sets, such that the speakers' actual experience of Australian English, as well as any stereotypes they held about its realization, likely differed considerably across our two data-sets.

Overall, our results suggest that Australian primes affect speech production. The exact manner in which this manifests is mediated by the salience and extent of the phonetic difference between dialects, and by individuals' own levels of experience of the dialects. These studies provide a combination of ecological validity and experimental control, and together reinforce our hypothesis: contextual activation of a known dialect can affect speech production patterns.

Finding that mention of Australia in an experimental context was able to shift speech productions suggests that such effects can occur in a fairly automatic manner. The fact that we get the same overall pattern in our corpus data suggests that what we see in the corpus is not (or not exclusively) identity-driven. This work therefore speaks to recent discussions in the literature about the degree to which topic-based variation in speech is automatic, or driven by identity (Love & Walker, 2013). Our results suggest that – while identity may play a role – it is unlikely to be the sole driver of the topic-based variation that is observed in the literature; as such variation can clearly arise through contextual priming effects. Moreover, these results fit well within the experimental accommodation literature that has observed shifts in speech productions in a social devoid context, which also supports the idea that accommodation occurs in an automatic fashion. Furthermore, it should also be stressed that the results of this study should be taken in a broader context, as it is unlikely that these effects are limited to NZE and AE speech, just as phonetic accommodation has been found cross-linguistically.

The observed shift of KIT in our corpus analysis is directly predicted by models containing experience-based representations with socio-indexical associations (e.g. Foulkes & Docherty, 2006). KIT saliently varies between Australian and New Zealand English. If some previously encountered Australian KIT vowels were associated with Australia, then contextual priming of Australia can shift productions towards these variants.

The observed shift in TRAP in our corpus analysis is also predicted by experience-based models with representations of socio-indexical associations (e.g. Foulkes & Docherty, 2006). Although TRAP is not a highly salient feature, it is known appropriately between the dialects. In order to elicit the appropriate speech change (AE-like), additional primes are needed in addition to the context.

In the class of models just described, we can also address why the degree of shift in the predicted direction is mediated by experience with Australia for DRESS. As mentioned, differences in DRESS appears to be a non-salient feature between the dialects. The location of NZ and AE DRESS, for example, is often confused and misattributed (Ludwig, 2007). As such, this may explain why those with ample previous experience of Australian speakers show a greater degree of shift in the predicted direction than those with lesser experience.

Problematic for current models, however, are the results showing that speakers with little experience of Australia actually shift their DRESS in the opposite direction (i.e. toward less Australian variants). It appears that such speakers are relying on a false stereotype of the realization of Australian DRESS. Previous results in the literature indicate that some NZers believe that Australians actually use a higher fronter DRESS vowel than they do (Ludwig, 2007). Our results can be interpreted as showing that, in an Australian-primed context, New Zealanders with less experience of Australia shift in the direction of the variant they *perceive* as being more Australian. This raises interesting questions about the nature of representations, of what is stored, and how beliefs and expectations can affect productions.

Acknowledgments

This work utilizes the Canterbury Corpus, which is part of the Origins of New Zealand English (ONZE) archives at the University of Canterbury. The data was collected by members of the NZ English class of the Linguistics Department, University of Canterbury. The work done by members of the ONZE project in preparing the data, making transcripts and obtaining background information is gratefully acknowledged. This work has been supported by a Royal Society Marsden Fund grant (M1123-UOC0809) and a Rutherford Discovery Fellowship (E5909) awarded to the second author, and by a University of Canterbury Summer Scholarship awarded to the third author. Helpful assistance has been provided by Morgana Mountfort-Davies and Alia Hope-Wilson. We also acknowledge the helpful feedback received from our colleagues at NZILBB, the audience of the New Zealand Linguistics Society Conference, our anonymous reviewers, and the editors of this special issue.

Appendix A. Search terms used to identify sections of speech in the corpus referring to Australia

Australia
Aussie
Oz
Koala
Kangaroo
Wallabies (Rugby Team)
Sydney
Brisbane
Perth
Melbourne
Canberra
Adelaide
Cairns
Hobart
Gold Coast

Great Barrier Reef
Queensland
New South Wales
Northern Territory
Aborigine

Appendix B. Critical word lists used in experiment

Neutral1	Neutral2	AUS1	AUS2	KIT	DRESS
Aubergine	Barb	Aborigine	Bondi beach	Bib	Bent
Marmalade	Drive	Adelaide	Boomerang	Finch	Drench
Boss	Game	Aussie	Dingo	Kink	Fend
Awesome	Harsh	Australia	Kookaburra	Pip	Fret
Brighten	Hawk	Brisbane	Oz	Pith	Hex
Cane	Hide	Cairns	Platypus	Skimp	Mend
Classroom	Merge	Canberra		Skit	Mesh
Glasshouse	Piece	Gold coast		Snip	Nets
Hopeful	Sleep	Hobart		Spit	Peck
Capable	Spike	Kangaroo		Stint	Peg
Cholera	Steak	Koala		Sync	Prep
Microphone	Stir	Melbourne		Tiff	Pressed
Outbreak	Twirl	Outback		Twit	Quest
Path	Waist	Perth		Whiz	Rex
Quartermaster	Walk	Queensland		Zing	Sled
Sardine		Sydney			Speck
Troublemaker		Tasmania			Specs
Wallflower		Wallabies			Spent
					Stead
					Stench
					Temp
					Trek
					Trench
					Vest
					Vet
					Vex
					Web
					Webbed
					Wed
					Wedge
					Weld
					Wench
					Whence
					Whet
					Wretch
					Zest

Appendix C. Post-experiment questionnaire regarding experience with Australian English (from Drager et al., 2010)

1. Have you ever been to Australia?
   No Yes
   – For how long total? _____
   – How long ago since you were last there? _____
2. Are you personally acquainted with many Australians?
   No Yes – Give relationship:_____
3. How many Australians would you come into personal contact with during an average week?
None 1–5 5–10 10+
4. Which of the following programmes (if any) do you watch?
Home and Away *Rove* *McLeod's Daughters* *Neighbours*
Blue Heelers *Australian News*

Kath and Kim

5. Which of the following sentences best describes you:

I never watch TV.
I sometimes watch TV.
I watch a whole lot of TV.

6. Do you follow international sport?

No Yes– Which Sports?

Appendix D. Post-experiment questionnaire regarding explicit Australian attitudes, from Drager et al. (2010)

Please read the following sentences and for each one, provide a rating indicating how much you agree with it. Please provide an answer for every example.

	Strongly disagree	Disagree	Agree	Strongly agree
1. Australians and New Zealanders are very similar.	1	2	3	4
2. I find it annoying when people get New Zealanders and Australians confused.	1	2	3	4
3. Australians and New Zealanders agree on the important issues.	1	2	3	4
4. Australia has a good human rights record.	1	2	3	4
5. In most sports, the team I most want New Zealand to beat is Australia.	1	2	3	4
6. Most stereotypes about Australians are false.	1	2	3	4
7. I would enjoy living in Australia.	1	2	3	4
8. Australia is a place I'd like to go for a holiday.	1	2	3	4
9. New Zealanders and Australians have very similar accents.	1	2	3	4
10. In most sports, if New Zealand is not playing, then I tend to support Australia.	1	2	3	4

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H. (2009). *Language R: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics"*. R package version 0.955.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*, 177–189.

Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society*, *39*(4), 437–456.

Bates, D. M. & Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-32.

Bauer, L. (1986). Notes on New Zealand English phonetics and phonology. *English World-Wide*, *7*(2), 225–258.

Bauer, L., Warren, P., Bardsley, D., Kennedy, M., & Major, G. (2007). *Illustrations of the IPA: New Zealand English Journal of the International Phonetic Association*, *37*(1), 97–102.

Becker, K. (2009). /r/ and the construction of place identity on New York City's Lower East Side. *Journal of Sociolinguistics*, *13*(5), 634–658.

Bell, A. (1984). Language style as audience design. *Language and Society*, *13*, 145–204.

Boersma, P., & Weenink, D. (2008). *Praat: Doing phonetics by computer* (Version 5.0.32) [Computer program]. Retrieved January 2008 from ⟨www.praat.org/⟩.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

Cox, F., & Palethorpe, S. (2008). Reversal of short front vowel raising in Australian English. In *Proceedings of interspeech* (pp. 342–345). Brisbane, Australia.

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, *64*, 145–173.

Drager, K., Hay, J., & Walker, A. (2010). Pronounced rivalries: Attitudes and speech production. *Te Reo*, *53*, 27–53.

Drager, K., & Hay, J. (2010). Visual subliminal primes affect vowel perception. Poster presented at the *12th conference on laboratory phonology*. Albuquerque, July 2010.

Eckert, P. (2000). *Linguistic variation as social practice*. Oxford, UK: Blackwell.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*, 409–438.

Fromont, R., & Hay, J. (2008). ONZE Miner: Development of a browser-based research tool. *Corpora*, *3*(2), 173–193.

Fromont, R., & Hay, J. (2012). LaBB-CAT: An annotation store. In *Proceedings of the Australasian language technology workshop* (*ALTA*) (Vol. 10, pp. 113–117). Dunedin, New Zealand: University of Otago, 4–6 December 2012. ⟨http://www.aclweb.org/anthology/U/U12/U12-2015⟩.

Giles, H., Coupland, J., & Coupland, N. (1991). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge: Cambridge University Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.

Goldinger, S. D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review*, *11*, 716–722.

Gordon, E., Maclagan, M. A., & Hay, J. (2007). The ONZE Corpus. In J. C. Beal, K. P. Corrigan, & H. Moisl (Eds.), *Models and methods in the handling of unconventional digital corpora: Volume 2, Diachronic corpora* (pp. 82–104). Basingstoke, Hampshire: Palgrave Macmillan.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*, 865–892.

Hay, J., Drager, K., & Thomas, B. (2013). Using nonsense words to investigate vowel merger. *English Language and Linguistics*, *17*(2), 241–269.

Hay, J., Drager, K., & Warren, P. (2009). Careful who you talk to: An effect of experimenter identity on the production of the near/square merger in New Zealand English. *Australian Journal of Linguistics*, *29*, 269–285.

Hay, J., Maclagan, M., & Gordon, M. (2008). *New Zealand English*. Edinburgh University Press.

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, *23*, 351–379.

Hay, J., Pierrehumbert, J., Walker, A., & LaShell, P. (2014). Tracking word frequency effects through 130 years of sound change (under review).

Hay, J., Warren, P., & Drager, K. (2010). Short-term exposure to one dialect affects processing of another. *Language and Speech*, *53*, 447–471.

Jannedy, S., Weirich, M., & Brunner, J. (2011). The effect of inferences on the perceptual categorization of Berlin German fricatives. In *Proceedings of the international congress of phonetic sciences (ICPhS)*. Hong Kong.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*, 125–156, http://dx.doi.org/10.1515/labphon.2011.004.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Lawson, R. (2009). *Sociolinguistic constructions of identity among adolescent males in Glasgow* (Doctoral dissertation). Scotland, UK: University of Glasgow.

Lewandowski, N. (2012). *Talent in nonnative phonetic convergence* (Dissertation). Universität Stuttgart.

Love, J., & Walker, A. (2013). Football versus football: Effect of topic on /r/ realization in American and English sports fans. *Language and Speech*, *56*, 443–460.
Ludwig, I. (2007). *Identification of New Zealand English and Australian English based on stereotypical accent markers* (Master's thesis). University of Canterbury.
Maclagan, M., & Hay, J. (2007). Getting fed up with our feet: Contrast maintenance and the New Zealand English "short" front vowel shift. *Language Variation and Change*, *19*, 1–25.
McKenzie, J. (2005). But he's not supposed to see me in my weeding dress! The relationship between DRESS and FLEECE in modern New Zealand English. *New Zealand English Journal*, *19*, 13–25.
Mendoza-Denton, N., Hay, J., & Jannedy, S. (2003). Probabilistic sociolinguistics: Beyond variable rules. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 97–139). Cambridge, MA: MIT Press.
Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2010). Alignment to visual speech information. *Attention, Perception, & Psychophysics*, *72*, 1614–1625.
Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2014). Is speech alignment to talkers or tasks?. *Attention, Perception, & Psychophysics*, http://dx.doi.org/10.3758/APP.72.6.1614, in press.
Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender difference in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*, *21*, 422–432.
Nguyen, N., Dufour, S., & Brunellière, A. (2012). Does imitation facilitate word recognition in a non-native regional accent?. *Frontiers in Psychology*, *3*, 480, http://dx.doi.org/10.3389/fpsyg.2012.00480.
Niedzielski, N. (1997). *The effect of social information on the phonetic perception of sociolinguistic variables* (Doctoral dissertation). Santa Barbara: University of California.
Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*, 62–85.
Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*, 132–142.
Olmstead, A., Viswanathan, N., Aivar, M. P., & Manuel, S. (2013). Comparison of native and non-native phone imitation by English and Spanish speakers. *Frontiers in Psychology*, *4*, 475, http://dx.doi.org/10.3389/fpsyg.2013.00475.
Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral & Brain Sciences*, *27*(2), 169–226.
Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee, & Paul J. Hopper (Eds.), *Frequency effects and emergent grammar* (pp. 137–158). Amsterdam & Philadelphia: John Benjamins.
Pierrehumbert, J. (2002). Word-specific phonetics. In Carlos Gussenhoven, & Natasha Warner (Eds.), *Laboratory Phonology 7* (pp. 101–139). Berlin & New York: Mouton de Gruyter.
Pierrehumbert, J. (2006). The next toolkit. *Journal of Phonetics*, *34*, 516–530.
R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ⟨http://www.R-project.org⟩.
Sanchez, K. (2011). *Do you hear what I see? The voice and face of a talker similarly influence the speech of multiple listeners* (Doctoral dissertation). Riverside: University of California.
Sanchez, K., Miller, R. M., & Rosenblum, L. D. (2010). Visual influences on alignment to voice onset time. *Journal of Speech, Language, and Hearing Research*, *53*, 262–272.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools, Inc.
Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics*, *66*, 422–429.
Tilsen, S. (2009). Subphonemic and cross-phonemic priming in vowel shadowing: Evidence for the involvement of exemplars in production. *Journal of Phonetics*, *37*, 276–296.
Watson, C., Harrington, J., & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics*, *18*, 185–207.
Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.